

SEQUENCE LISTING

(1) GENERAL INFORMATION:

5

(i) APPLICANT: Tryggvason, Karl
Kestila, Marjo
Lenkkeri, Ulla
Mannikko, Minna

10

(ii) TITLE OF INVENTION: Neph rin Gene and Protein

(iii) NUMBER OF SEQUENCES: 6

15

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
(B) STREET: 300 S. Wacker Drive, Suite 3200
(C) CITY: Chicago
(D) STATE: IL
(E) COUNTRY: USA
(F) ZIP: 60606

20

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

30

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Chao, Mark
(B) REGISTRATION NUMBER: 37,293
(C) REFERENCE/DOCKET NUMBER: 97,842

35

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (312)913-0001
(B) TELEFAX: (312)913-0002

40

(2) INFORMATION FOR SEQ ID NO:1:

45

(i) SEQUENCE CHARACTERISTICS:

50

(A) LENGTH: 4285 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: not relevant

55

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
 (A) NAME/KEY: sig_peptide
 (B) LOCATION: 1..66
 (D) OTHER INFORMATION: /note= "putative signal peptide"

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..3723

(ix) FEATURE:
 (A) NAME/KEY: mat_peptide
 (B) LOCATION: 67..3723

(ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION: 121..122
 (D) OTHER INFORMATION: /note= "deletion mutation
 FIN-Major"

(ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION: 3800..3804
 (D) OTHER INFORMATION: /note= "nonsense mutation in exon
 26 FIN-Minor"

(ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION: 3178..3258
 (D) OTHER INFORMATION: /note= "putative transmembrane
 domain"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

35	ATG GCC CTG GGG ACG ACG CTC AGG GCT TCT CTC CTG CTC CTG GGG CTG	48
	Met Ala Leu Gly Thr Thr Leu Arg Ala Ser Leu Leu Leu Gly Leu	
	-22 -20 -15 -10	
40	CTG ACT GAA GGC CTG GCG CAG TTG GCG ATT CCT GCC TCC GTT CCC CGG	96
	Leu Thr Glu Gly Leu Ala Gln Leu Ala Ile Pro Ala Ser Val Pro Arg	
	-5 1 5 10	
45	GGC TTC TGG GCC CTG CCT GAA AAC CTG ACG GTG GTG GAG GGG GCC TCA	144
	Gly Phe Trp Ala Leu Pro Glu Asn Leu Thr Val Val Glu Gly Ala Ser	
	15 20 25	
50	GTG GAG CTG CGT TGT GGG GTC AGC ACC CCT GGC AGT GCG GTG CAA TGG	192
	Val Glu Leu Arg Cys Gly Val Ser Thr Pro Gly Ser Ala Val Gln Trp	
	30 35 40	
55	GCC AAA GAT GGG CTG CTC CTG GGC CCC GAC CCC AGG ATC CCA GGC TTC	240
	Ala Lys Asp Gly Leu Leu Leu Gly Pro Asp Pro Arg Ile Pro Gly Phe	
	45 50 55	
	CCG AGG TAC CGC CTG GAA GGG GAC CCT GCT AGA GGT GAA TTC CAC CTG	288
	Pro Arg Tyr Arg Leu Glu Gly Asp Pro Ala Arg Gly Glu Phe His Leu	

	60	65	70	
	CAC ATC GAG GCC TGT GAC CTC AGC GAT GAC GCG GAG TAT GAG TGC CAG			336
5	His Ile Glu Ala Cys Asp Leu Ser Asp Asp Ala Glu Tyr Glu Cys Gln	80	85	90
	75			
	GTC GGC CGC TCT GAG ATG GGG CCC GAG CTC GTG TCT CCC AGA GTG ATC			384
	Val Gly Arg Ser Glu Met Gly Pro Glu Leu Val Ser Pro Arg Val Ile	95	100	105
10				
	CTC TCC ATC CTG GTT CCT CCC AAG CTG CTC CTG CTG ACC CCA GAG GCA			432
	Leu Ser Ile Leu Val Pro Pro Lys Leu Leu Leu Thr Pro Glu Ala	110	115	120
15	GGC ACC ATG GTC ACC TGG GTA GCT GGG CAG GAG TAC GTG GTC AAC TGT			480
	Gly Thr Met Val Thr Trp Val Ala Gly Gln Glu Tyr Val Val Asn Cys	125	130	135
20	GTG TCT GGG GAC GCG AAG CCA GCA CCT GAC ATC ACC ATT CTC CTG AGT			528
	Val Ser Gly Asp Ala Lys Pro Ala Pro Asp Ile Thr Ile Leu Leu Ser	140	145	150
25	GGA CAG ACA ATA TCT GAC ATC TCT GCA AAC GTG AAC GAG GGC TCC CAG			576
	Gly Gln Thr Ile Ser Asp Ile Ser Ala Asn Val Asn Glu Gly Ser Gln	155	160	165
	CAG AAA CTC TTC ACT GTG GAG GCC ACA GCC AGG GTG ACA CCC CGG AGC			624
	Gln Lys Leu Phe Thr Val Glu Ala Thr Ala Arg Val Thr Pro Arg Ser	175	180	185
30				
	TCA GAT AAT AGG CAG TTG CTG GTC TGT GAG GCG TCT AGC CCA GCA CTG			672
	Ser Asp Asn Arg Gln Leu Leu Val Cys Glu Ala Ser Ser Pro Ala Leu	190	195	200
35	GAG GCC CCC ATC AAG GCC TCA TTC ACC GTG AAT GTT CTG TTC CCT CCA			720
	Glu Ala Pro Ile Lys Ala Ser Phe Thr Val Asn Val Leu Phe Pro Pro	205	210	215
40	GGA CCC CCT GTC ATC GAG TGG CCA GGC CTG GAT GAG GGG CAC GTG CGG			768
	Gly Pro Pro Val Ile Glu Trp Pro Gly Leu Asp Glu Gly His Val Arg	220	225	230
45	GCA GGA CAG AGC TTG GAG CTG CCG TGC GTG GCC CGA GGG GGT AAT CCC			816
	Ala Gly Gln Ser Leu Glu Leu Pro Cys Val Ala Arg Gly Gly Asn Pro	235	240	245
	TTA GCC ACA CTG CAG TGG CTG AAG AAT GGC CAG CCG GTG TCC ACA GCG			864
	Leu Ala Thr Leu Gln Trp Leu Lys Asn Gly Gln Pro Val Ser Thr Ala	255	260	265
50				
	TGG GGC ACA GAG CAC ACC CAG GCG GTG GCC CGC AGT GTG CTG GTG ATG			912
	Trp Gly Thr Glu His Thr Gln Ala Val Ala Arg Ser Val Leu Val Met	270	275	280
55	ACC GTG AGG CCA GAA GAC CAT GGA GCG CAG CTC AGC TGC GAG GCC CAC			960
	Thr Val Arg Pro Glu Asp His Gly Ala Gln Leu Ser Cys Glu Ala His	285	290	295

	AAC AGC GTG TCT GCA GGG ACC CAG GAG CAC GGC ATC ACA CTG CAG GTC Asn Ser Val Ser Ala Gly Thr Gln Glu His Gly Ile Thr Leu Gln Val 300 305 310	1008
5	ACC TTT CCC CCT AGT GCC ATT ATT ATC TTG GGA TCT GCA TCC CAG ACT Thr Phe Pro Pro Ser Ala Ile Ile Ile Leu Gly Ser Ala Ser Gln Thr 315 320 325 330	1056
10	GAG AAC AAG AAC GTG ACA CTC TCC TGT GTC AGC AAG TCC AGT CGC CCG Glu Asn Lys Asn Val Thr Leu Ser Cys Val Ser Lys Ser Ser Arg Pro 335 340 345	1104
15	CGG GTT CTG CTA CGA TGG TGG CTG GGC TGG CGG CAG CTG CTG CCC ATG Arg Val Leu Leu Arg Trp Trp Leu Gly Trp Arg Gln Leu Leu Pro Met 350 355 360	1152
	GAG GAG ACA GTC ATG GAT GGA CTG CAT GGC GGT CAC ATC TCC ATG TCC Glu Glu Thr Val Met Asp Gly Leu His Gly Gly His Ile Ser Met Ser 365 370 375	1200
20	AAC CTG ACA TTC CTG GCG CGG CGG GAG GAC AAC GGT CTG ACC CTC ACA Asn Leu Thr Phe Leu Ala Arg Arg Glu Asp Asn Gly Leu Thr Leu Thr 380 385 390	1248
25	TGT GAG GCC TTC AGT GAA GCC TTC ACC AAG GAG ACC TTC AAG AAG TCG Cys Glu Ala Phe Ser Glu Ala Phe Thr Lys Glu Thr Phe Lys Lys Ser 395 400 405 410	1296
30	CTC ATC CTG AAC GTA AAA TAT CCC GCC CAG AAA CTG TGG ATT GAG GGT Leu Ile Leu Asn Val Lys Tyr Pro Ala Gln Lys Leu Trp Ile Glu Gly 415 420 425	1344
35	CCC CCA GAG GGC CAG AAG CTC CGG GCT GGG ACC CGG GTG AGG CTG GTG Pro Pro Glu Gly Gln Lys Leu Arg Ala Gly Thr Arg Val Arg Leu Val 430 435 440	1392
	TGT TTG GCT ATC GGG GGC AAC CCA GAG CCC TCC CTC ATG TGG TAC AAG Cys Leu Ala Ile Gly Gly Asn Pro Glu Pro Ser Leu Met Trp Tyr Lys 445 450 455	1440
40	GAC TCG CGC ACC GTG ACC GAG TCG CGG CTG CCG CAG GAG TCG CGG CGC Asp Ser Arg Thr Val Thr Glu Ser Arg Leu Pro Gln Glu Ser Arg Arg 460 465 470	1488
45	GTG CAT CTC GGC AGC GTG GAG AAA TCT GGG AGC ACC TTC TCC CGA GAG Val His Leu Gly Ser Val Glu Lys Ser Gly Ser Thr Phe Ser Arg Glu 475 480 485 490	1536
50	CTG GTG CTG GTC ACA GGG CCG TCG GAC AAC CAG GCC AAG TTC ACG TGC Leu Val Leu Val Thr Gly Pro Ser Asp Asn Gln Ala Lys Phe Thr Cys 495 500 505	1584
55	AAG GCT GGA CAG CTC AGC GCG TCC ACG CAG CTG GCG GTG CAG TTT CCC Lys Ala Gly Gln Leu Ser Ala Ser Thr Gln Leu Ala Val Gln Phe Pro 510 515 520	1632

	CCA ACT AAC GTG ACG ATC CTG GCC AAC GCA TCC GCA CTG CGC CCG GGA	1680
	Pro Thr Asn Val Thr Ile Leu Ala Asn Ala Ser Ala Leu Arg Pro Gly	
	525 530 535	
5	GAC GCC TTA AAC TTG ACA TGC GTC AGC GTC AGC AGC AAT CCG CCG GTC	1728
	Asp Ala Leu Asn Leu Thr Cys Val Ser Val Ser Ser Asn Pro Pro Val	
	540 545 550	
10	AAC TTG TCC TGG GAC AAG GAA GGG GAG AGG CTG GAG GGC GTG GCC GCC	1776
	Asn Leu Ser Trp Asp Lys Glu Gly Glu Arg Leu Glu Gly Val Ala Ala	
	555 560 565 570	
15	CCA CCC CGG AGA GCC CCA TTC AAA GGC TCC GCC GCC GCC AGG AGC GTC	1824
	Pro Pro Arg Arg Ala Pro Phe Lys Gly Ser Ala Ala Ala Arg Ser Val	
	575 580 585	
20	CTT CTG CAA GTG TCA TCC CGC GAT CAT GGC CAG CGC GTG ACC TGC CGC	1872
	Leu Leu Gln Val Ser Ser Arg Asp His Gly Gln Arg Val Thr Cys Arg	
	590 595 600	
25	GCC CAC AGC GCC GAG CTC CGC GAA ACC GTG AGC TCC TTC TAT CGC CTC	1920
	Ala His Ser Ala Glu Leu Arg Glu Thr Val Ser Ser Phe Tyr Arg Leu	
	605 610 615	
30	AAC GTA CTG TAC CGT CCA GAG TTC CTG GGG GAG CAG GTG CTG GTG GTG	1968
	Asn Val Leu Tyr Arg Pro Glu Phe Leu Gly Glu Gln Val Leu Val Val	
	620 625 630	
35	ACC GCG GTG GAG CAG GGC GAG GCG TTG CTG CCC GTG TCC GTG TCC GCT	2016
	Thr Ala Val Glu Gln Gly Glu Ala Leu Leu Pro Val Ser Val Ser Ala	
	635 640 645 650	
40	AAC CCC GCC CCC GAG GCC TTC AAC TGG ACC TTC CGC GGC TAT CGC CTC	2064
	Asn Pro Ala Pro Glu Ala Phe Asn Trp Thr Phe Arg Gly Tyr Arg Leu	
	655 660 665	
45	AGT CCA GCG GGC GGC CCC CGG CAT CGC ATC CTG TCC AGC GGG GCT CTG	2112
	Ser Pro Ala Gly Pro Arg His Arg Ile Leu Ser Ser Gly Ala Leu	
	670 675 680	
50	CAT CTG TGG AAT GTG ACC CGC GCG GAC GAC GGC CTC TAT CAG CTG CAC	2160
	His Leu Trp Asn Val Thr Arg Ala Asp Asp Gly Leu Tyr Gln Leu His	
	685 690 695	
55	TGC CAG AAC TCT GAG GGC ACC GCG GAA GCG CGG CTG CGG CTG GAC GTG	2208
	Cys Gln Asn Ser Glu Gly Thr Ala Glu Ala Arg Leu Arg Leu Asp Val	
	700 705 710	
60	CAC TAT GCT CCC ACC ATC CGT GCC CTC CAG GAC CCC ACT GAG GTG AAC	2256
	His Tyr Ala Pro Thr Ile Arg Ala Leu Gln Asp Pro Thr Glu Val Asn	
	715 720 725 730	
65	GTC GGG GGT TCT GTG GAC ATA GTC TGC ACT GTC GAT GCC AAT CCC ATC	2304
	Val Gly Gly Ser Val Asp Ile Val Cys Thr Val Asp Ala Asn Pro Ile	
	735 740 745	

	CTC	CCG	GGC	ATG	TTC	AAC	TGG	GAG	AGA	CTG	GGA	GAT	GAG	GAG	GAC	2352	
	Leu	Pro	Gly	Met	Phe	Asn	Trp	Glu	Arg	Leu	Gly	Glu	Asp	Glu	Glu	Asp	
				750					755					760			
5	CAG	AGC	CTG	GAT	GAC	ATG	GAG	AAG	ATA	TCC	AGG	GGA	CCA	ACG	GGG	CGC	2400
	Gln	Ser	Leu	Asp	Asp	Met	Glu	Lys	Ile	Ser	Arg	Gly	Pro	Thr	Gly	Arg	
			765					770					775				
10	CTG	CGG	ATT	CAC	CAT	GCC	AAA	CTG	GCC	CAG	GCT	GGC	GCT	TAC	CAG	TGC	2448
	Leu	Arg	Ile	His	His	Ala	Lys	Leu	Ala	Gln	Ala	Gly	Ala	Tyr	Gln	Cys	
			780				785					790					
15	ATT	GTG	GAC	AAT	GGG	GTG	GCG	CCT	CCA	GCA	CGA	CGG	CTG	CTC	CGT	CTT	2496
	Ile	Val	Asp	Asn	Gly	Val	Ala	Pro	Pro	Ala	Arg	Arg	Leu	Leu	Arg	Leu	
			795			800					805					810	
20	GTT	GTC	AGA	TTT	GCC	CCC	CAG	GTG	GAG	CAC	CCC	ACT	CCC	CTA	ACT	AAG	2544
	Val	Val	Arg	Phe	Ala	Pro	Gln	Val	Glu	His	Pro	Thr	Pro	Leu	Thr	Lys	
				815						820					825		
25	GTG	GCT	GCA	GCT	GGA	GAC	AGC	ACC	AGT	TCT	GCC	ACC	CTC	CAC	TGC	CGT	2592
	Val	Ala	Ala	Ala	Gly	Asp	Ser	Thr	Ser	Ser	Ala	Thr	Leu	His	Cys	Arg	
				830					835					840			
30	GCC	CGA	GGT	GTC	CCC	AAC	ATC	GTT	TTC	ACT	TGG	ACA	AAA	AAC	GGG	GTC	2640
	Ala	Arg	Gly	Val	Pro	Asn	Ile	Val	Phe	Thr	Trp	Thr	Lys	Asn	Gly	Val	
			845					850					855				
35	CCT	CTG	GAT	CTC	CAA	GAT	CCC	AGG	TAC	ACG	GAG	CAC	ACA	TAC	CAC	CAG	2688
	Pro	Leu	Asp	Leu	Gln	Asp	Pro	Arg	Tyr	Thr	Glu	His	Thr	Tyr	His	Gln	
			860				865					870					
40	GGT	GGT	GTC	CAC	AGC	AGC	CTC	CTG	ACC	ATT	GCC	AAC	GTG	TCT	GCC	GCC	2736
	Gly	Gly	Val	His	Ser	Ser	Leu	Leu	Thr	Ile	Ala	Asn	Val	Ser	Ala	Ala	
			875			880					885					890	
45	CAG	GAT	TAC	GCC	CTC	TTC	ACA	TGT	ACA	GCC	ACC	AAC	GCC	CTT	GGC	TCG	2784
	Gln	Asp	Tyr	Ala	Leu	Phe	Thr	Cys	Thr	Ala	Thr	Asn	Ala	Leu	Gly	Ser	
				895						900					905		
50	GAC	CAA	ACC	AAC	ATT	CAA	CTT	GTC	AGC	ATC	AGC	CGC	CCT	GAC	CCT	CCA	2832
	Asp	Gln	Thr	Asn	Ile	Gln	Leu	Val	Ser	Ile	Ser	Arg	Pro	Asp	Pro	Pro	
				910					915					920			
55	TCA	GGA	TTA	AAG	GTT	GTG	AGT	CTG	ACC	CCA	CAC	TCC	GTG	GGG	CTG	GAG	2880
	Ser	Gly	Leu	Lys	Val	Val	Ser	Leu	Thr	Pro	His	Ser	Val	Gly	Leu	Glu	
			925					930					935				
60	TGG	AAG	CCT	GGC	TTT	GAT	GGG	GGC	CTG	CCA	CAG	AGG	TTC	TGC	ATC	AGG	2928
	Trp	Lys	Pro	Gly	Phe	Asp	Gly	Gly	Leu	Pro	Gln	Arg	Phe	Cys	Ile	Arg	

	CCC CAG GCC ACC ACC TTC ACG CTG ACT GGT CTA CAG CCT TCT ACA AGA	3024
	Pro Gln Ala Thr Thr Phe Thr Leu Thr Gly Leu Gln Pro Ser Thr Arg	
	975 980 985	
5	TAC AGG GTC TGG CTG CTG GCC AGT AAT GCC TTG GGG GAC AGT GGA CTG	3072
	Tyr Arg Val Trp Leu Leu Ala Ser Asn Ala Leu Gly Asp Ser Gly Leu	
	990 995 1000	
10	GCT GAC AAA GGG ACC CAG CTT CCC ATC ACT ACC CCA GGT CTC CAC CAG	3120
	Ala Asp Lys Gly Thr Gln Leu Pro Ile Thr Thr Pro Gly Leu His Gln	
	1005 1010 1015	
15	CCT TCT GGA GAA CCT GAA GAC CAG CTG CCC ACA GAG CCA CCT TCA GGA	3168
	Pro Ser Gly Glu Pro Glu Asp Gln Leu Pro Thr Glu Pro Pro Ser Gly	
	1020 1025 1030	
20	CCC TCG GGG CTG CCC CTG CTG CCT GTG CTG TTC GCT CTT GGG GGG CTT	3216
	Pro Ser Gly Leu Pro Leu Leu Pro Val Leu Phe Ala Leu Gly Gly Leu	
	1035 1040 1045 1050	
25	CTG CTC CTC TCC AAT GCC TCC TGT GTC GGG GGG GTC CTC TGG CAG CGG	3264
	Leu Leu Leu Ser Asn Ala Ser Cys Val Gly Gly Val Leu Trp Gln Arg	
	1055 1060 1065	
30	AGA CTC AGG CGT CTT GCT GAG GGC ATC TCA GAG AAG ACA GAG GCA GGG	3312
	Arg Leu Arg Arg Leu Ala Glu Gly Ile Ser Glu Lys Thr Glu Ala Gly	
	1070 1075 1080	
35	TCG GAA GAG GAC CGA GTC AGG AAC GAA TAT GAG GAG AGC CAG TGG ACA	3360
	Ser Glu Glu Asp Arg Val Arg Asn Glu Tyr Glu Glu Ser Gln Trp Thr	
	1085 1090 1095	
40	GGA GAG CGG GAC ACT CAG AGC TCC ACG GTC AGC ACA ACA GAG GCA GAG	3408
	Gly Glu Arg Asp Thr Gln Ser Ser Thr Val Ser Thr Thr Glu Ala Glu	
	1100 1105 1110	
45	CCG TAT TAC CGC TCC CTG AGG GAC TTC AGC CCC CAG CTG CCC CCG ACG	3456
	Pro Tyr Tyr Arg Ser Leu Arg Asp Phe Ser Pro Gln Leu Pro Pro Thr	
	1115 1120 1125 1130	
50	CAG GAG GAG GTG TCT TAT TCC CGA GGT TTC ACA GGT GAA GAT GAG GAT	3504
	Gln Glu Glu Val Ser Tyr Ser Arg Gly Phe Thr Thr Gly Glu Asp Glu Asp	
	1135 1140 1145	
55	ATG GCC TTC CCT GGG CAC TTG TAT GAT GAG GTA GAA AGA ACG TAC CCC	3552
	Met Ala Phe Pro Gly His Leu Tyr Asp Glu Val Glu Arg Thr Tyr Pro	
	1150 1155 1160	
60	CCG TCT GGA GCC TGG GGA CCC CTC TAC GAT GAA GTG CAG ATG GGA CCC	3600
	Pro Ser Gly Ala Trp Gly Pro Leu Tyr Asp Glu Val Gln Met Gly Pro	
	1165 1170 1175	
65	TGG GAC CTC CAC TGG CCT GAA GAC ACA TAT CAG GAT CCA AGA GGA ATC	3648
	Trp Asp Leu His Trp Pro Glu Asp Thr Tyr Gln Asp Pro Arg Gly Ile	
	1180 1185 1190	

TAT GAC CAG GTG GCC GGA GAC TTG GAC ACT CTG GAA CCC GAT TCT CTG 3696
 Tyr Asp Gln Val Ala Gly Asp Leu Asp Thr Leu Glu Pro Asp Ser Leu
 1195 1200 1205 1210

5 CCC TTC GAG CTG AGG GGA CAT CTG GTG TAAGAGCCCT CTCAACCCCA 3743
 Pro Phe Glu Leu Arg Gly His Leu Val
 1215

10 TTGTCCTGCA CCTGCAGGAA TTTACACTCC ACTGGTCTCT CTCATTACAG CCTGGGCCGA 3803
 GCTGGTTAGG TGAGCTCCAT AAAACCCAAA GGGACTTGGT GTCAGGAGAG GACATGGAGG 3863

GGGCTGAGTG ACAGAGATGG TTCAGCTGGT ACCAGAGTAG AAACAAGGTG CATCCTGGGG 3923

15 TTGGCTTTAG AAATAAACT TCTCCAAAAG GACAGGGCAG ATTGTAAACG TCGTCTCAAA 3983
 AATGAAATGC TGCCGGGTGC GGTGACTCAC GCCTATAATC CCAGCACTTT GGGAGGCTGA 4043

GGCGGGTGGA TCACCTGAGG TCAGGAGTTC GAGACCAGCC TGGCCAACAT GGTAAAACTC 4103

20 CATTCTACT AAAAATATAA AAAATTAGCC AGGAGTAGTG GCGCATGCCT GTAGTCCCAG 4163

CTACTTGGGA GGCTGATGCA TGAGAATTGC TTGAACCCAG GAGGCGGAGG TTGCAGTGAG 4223

25 CTGAGATCAC GCCACTGCAC TCCAGCCTGG GCGACAGAGC GAGATTCTGT CTCAAAAAAT 4283

AA 4285

30 (2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1241 amino acids

(B) TYPE: amino acid

35 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

40 Met Ala Leu Gly Thr Thr Leu Arg Ala Ser Leu Leu Leu Leu Gly Leu
 -22 -20 -15 -10

45 Leu Thr Glu Gly Leu Ala Gln Leu Ala Ile Pro Ala Ser Val Pro Arg
 -5 1 5 10

Gly Phe Trp Ala Leu Pro Glu Asn Leu Thr Val Val Glu Gly Ala Ser
 15 20 25

50 Val Glu Leu Arg Cys Gly Val Ser Thr Pro Gly Ser Ala Val Gln Trp
 30 35 40

Ala Lys Asp Gly Leu Leu Leu Gly Pro Asp Pro Arg Ile Pro Gly Phe
 45 50 55

55 Pro Arg Tyr Arg Leu Glu Gly Asp Pro Ala Arg Gly Glu Phe His Leu
 60 65 70

His Ile Glu Ala Cys Asp Leu Ser Asp Asp Ala Glu Tyr Glu Cys Gln
 75 80 85 90
 5 Val Gly Arg Ser Glu Met Gly Pro Glu Leu Val Ser Pro Arg Val Ile
 95 100 105
 Leu Ser Ile Leu Val Pro Pro Lys Leu Leu Leu Thr Pro Glu Ala
 110 115 120
 10 Gly Thr Met Val Thr Trp Val Ala Gly Gln Glu Tyr Val Val Asn Cys
 125 130 135
 Val Ser Gly Asp Ala Lys Pro Ala Pro Asp Ile Thr Ile Leu Leu Ser
 140 145 150
 15 Gly Gln Thr Ile Ser Asp Ile Ser Ala Asn Val Asn Glu Gly Ser Gln
 155 160 165 170
 20 Gln Lys Leu Phe Thr Val Glu Ala Thr Ala Arg Val Thr Pro Arg Ser
 175 180 185
 Ser Asp Asn Arg Gln Leu Leu Val Cys Glu Ala Ser Ser Pro Ala Leu
 190 195 200
 25 Glu Ala Pro Ile Lys Ala Ser Phe Thr Val Asn Val Leu Phe Pro Pro
 205 210 215
 Gly Pro Pro Val Ile Glu Trp Pro Gly Leu Asp Glu Gly His Val Arg
 220 225 230
 30 Ala Gly Gln Ser Leu Glu Leu Pro Cys Val Ala Arg Gly Gly Asn Pro
 235 240 245 250
 35 Leu Ala Thr Leu Gln Trp Leu Lys Asn Gly Gln Pro Val Ser Thr Ala
 255 260 265
 Trp Gly Thr Glu His Thr Gln Ala Val Ala Arg Ser Val Leu Val Met
 270 275 280
 40 Thr Val Arg Pro Glu Asp His Gly Ala Gln Leu Ser Cys Glu Ala His
 285 290 295
 Asn Ser Val Ser Ala Gly Thr Gln Glu His Gly Ile Thr Leu Gln Val
 300 305 310
 45 Thr Phe Pro Pro Ser Ala Ile Ile Ile Leu Gly Ser Ala Ser Gln Thr
 315 320 325 330
 50 Glu Asn Lys Asn Val Thr Leu Ser Cys Val Ser Lys Ser Ser Arg Pro
 335 340 345
 Arg Val Leu Leu Arg Trp Trp Leu Gly Trp Arg Gln Leu Leu Pro Met
 350 355 360
 55 Glu Glu Thr Val Met Asp Gly Leu His Gly Gly His Ile Ser Met Ser
 365 370 375

Asn Leu Thr Phe Leu Ala Arg Arg Glu Asp Asn Gly Leu Thr Leu Thr
 380 385 390

5 Cys Glu Ala Phe Ser Glu Ala Phe Thr Lys Glu Thr Phe Lys Lys Ser
 395 400 405 410

Leu Ile Leu Asn Val Lys Tyr Pro Ala Gln Lys Leu Trp Ile Glu Gly
 415 420 425

10 Pro Pro Glu Gly Gln Lys Leu Arg Ala Gly Thr Arg Val Arg Leu Val
 430 435 440

15 Cys Leu Ala Ile Gly Gly Asn Pro Glu Pro Ser Leu Met Trp Tyr Lys
 445 450 455

Asp Ser Arg Thr Val Thr Glu Ser Arg Leu Pro Gln Glu Ser Arg Arg
 460 465 470

20 Val His Leu Gly Ser Val Glu Lys Ser Gly Ser Thr Phe Ser Arg Glu
 475 480 485 490

Leu Val Leu Val Thr Gly Pro Ser Asp Asn Gln Ala Lys Phe Thr Cys
 495 500 505

25 Lys Ala Gly Gln Leu Ser Ala Ser Thr Gln Leu Ala Val Gln Phe Pro
 510 515 520

30 Pro Thr Asn Val Thr Ile Leu Ala Asn Ala Ser Ala Leu Arg Pro Gly
 525 530 535

Asp Ala Leu Asn Leu Thr Cys Val Ser Val Ser Ser Asn Pro Pro Val
 540 545 550

35 Asn Leu Ser Trp Asp Lys Glu Gly Glu Arg Leu Glu Gly Val Ala Ala
 555 560 565 570

Pro Pro Arg Arg Ala Pro Phe Lys Gly Ser Ala Ala Ala Arg Ser Val
 575 580 585

40 Leu Leu Gln Val Ser Ser Arg Asp His Gly Gln Arg Val Thr Cys Arg
 590 595 600

45 Ala His Ser Ala Glu Leu Arg Glu Thr Val Ser Ser Phe Tyr Arg Leu
 605 610 615

Asn Val Leu Tyr Arg Pro Glu Phe Leu Gly Glu Gln Val Leu Val Val
 620 625 630

50 Thr Ala Val Glu Gln Gly Glu Ala Leu Leu Pro Val Ser Val Ser Ala
 635 640 645 650

Asn Pro Ala Pro Glu Ala Phe Asn Trp Thr Phe Arg Gly Tyr Arg Leu
 655 660 665

55 Ser Pro Ala Gly Gly Pro Arg His Arg Ile Leu Ser Ser Gly Ala Leu
 670 675 680

His Leu Trp Asn Val Thr Arg Ala Asp Asp Gly Leu Tyr Gln Leu His
 685 690 695
 5 Cys Gln Asn Ser Glu Gly Thr Ala Glu Ala Arg Leu Arg Leu Asp Val
 700 705 710
 His Tyr Ala Pro Thr Ile Arg Ala Leu Gln Asp Pro Thr Glu Val Asn
 715 720 725 730
 10 Val Gly Gly Ser Val Asp Ile Val Cys Thr Val Asp Ala Asn Pro Ile
 735 740 745
 Leu Pro Gly Met Phe Asn Trp Glu Arg Leu Gly Glu Asp Glu Glu Asp
 15 750 755 760
 Gln Ser Leu Asp Asp Met Glu Lys Ile Ser Arg Gly Pro Thr Gly Arg
 765 770 775
 20 Leu Arg Ile His His Ala Lys Leu Ala Gln Ala Gly Ala Tyr Gln Cys
 780 785 790
 Ile Val Asp Asn Gly Val Ala Pro Pro Ala Arg Arg Leu Leu Arg Leu
 795 800 805 810
 25 Val Val Arg Phe Ala Pro Gln Val Glu His Pro Thr Pro Leu Thr Lys
 815 820 825
 Val Ala Ala Ala Gly Asp Ser Thr Ser Ser Ala Thr Leu His Cys Arg
 30 830 835 840
 Ala Arg Gly Val Pro Asn Ile Val Phe Thr Trp Thr Lys Asn Gly Val
 845 850 855
 35 Pro Leu Asp Leu Gln Asp Pro Arg Tyr Thr Glu His Thr Tyr His Gln
 860 865 870
 Gly Gly Val His Ser Ser Leu Leu Thr Ile Ala Asn Val Ser Ala Ala
 875 880 885 890
 40 Gln Asp Tyr Ala Leu Phe Thr Cys Thr Ala Thr Asn Ala Leu Gly Ser
 895 900 905
 Asp Gln Thr Asn Ile Gln Leu Val Ser Ile Ser Arg Pro Asp Pro Pro
 45 910 915 920
 Ser Gly Leu Lys Val Val Ser Leu Thr Pro His Ser Val Gly Leu Glu
 925 930 935
 50 Trp Lys Pro Gly Phe Asp Gly Gly Leu Pro Gln Arg Phe Cys Ile Arg
 940 945 950
 Tyr Glu Ala Leu Gly Thr Pro Gly Phe His Tyr Val Asp Val Val Pro
 955 960 965 970
 55 Pro Gln Ala Thr Thr Phe Thr Leu Thr Gly Leu Gln Pro Ser Thr Arg
 975 980 985

Tyr Arg Val Trp Leu Leu Ala Ser Asn Ala Leu Gly Asp Ser Gly Leu
 990 995 1000
 5 Ala Asp Lys Gly Thr Gln Leu Pro Ile Thr Thr Pro Gly Leu His Gln
 1005 1010 1015
 Pro Ser Gly Glu Pro Glu Asp Gln Leu Pro Thr Glu Pro Pro Ser Gly
 1020 1025 1030
 10 Pro Ser Gly Leu Pro Leu Leu Pro Val Leu Phe Ala Leu Gly Gly Leu
 1035 1040 1045 1050
 Leu Leu Leu Ser Asn Ala Ser Cys Val Gly Gly Val Leu Trp Gln Arg
 15 1055 1060 1065
 Arg Leu Arg Arg Leu Ala Glu Gly Ile Ser Glu Lys Thr Glu Ala Gly
 1070 1075 1080
 20 Ser Glu Glu Asp Arg Val Arg Asn Glu Tyr Glu Glu Ser Gln Trp Thr
 1085 1090 1095
 Gly Glu Arg Asp Thr Gln Ser Ser Thr Val Ser Thr Thr Glu Ala Glu
 1100 1105 1110
 25 Pro Tyr Tyr Arg Ser Leu Arg Asp Phe Ser Pro Gln Leu Pro Pro Thr
 1115 1120 1125 1130
 Gln Glu Glu Val Ser Tyr Ser Arg Gly Phe Thr Gly Glu Asp Glu Asp
 30 1135 1140 1145
 Met Ala Phe Pro Gly His Leu Tyr Asp Glu Val Glu Arg Thr Tyr Pro
 1150 1155 1160
 35 Pro Ser Gly Ala Trp Gly Pro Leu Tyr Asp Glu Val Gln Met Gly Pro
 1165 1170 1175
 Trp Asp Leu His Trp Pro Glu Asp Thr Tyr Gln Asp Pro Arg Gly Ile
 1180 1185 1190
 40 Tyr Asp Gln Val Ala Gly Asp Leu Asp Thr Leu Glu Pro Asp Ser Leu
 1195 1200 1205 1210
 Pro Phe Glu Leu Arg Gly His Leu Val
 45 1215

(2) INFORMATION FOR SEQ ID NO:3:

- 50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 55 (ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "primer exon 2 5'UTR"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

5 GAGAAAGCCA GACAGACGCA G

21

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "primer intron 2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

20 AGCTTCCGCT GGTGGCT

17

(2) INFORMATION FOR SEQ ID NO:5:

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "primer intron 23"

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTCGGGGAGA CCCACCC

17

(2) INFORMATION FOR SEQ ID NO:6:

40 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

45 (ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "primer intron 26"

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

55 CCTGATGCTA ACGGCAGGGC

20